CHAPTER 2

BIVARIATE PARAMETRIC STATISTICAL TESTS OF HYPOTHESES FOR MATCHING SIMILARITY OF IMAGES

2.1 INTRODUCTION

Representation is fundamental in both biological and computational vision system [Jenk97]. An effective, efficient and suitable representation is the key starting point for building image processing and computer vision system. In many ways, the success or failure of an algorithm depends greatly on an appropriately designed representation. In the computer vision community, it is a common practice to classify representation schemes as low-level, intermediate-level and high-level. Low-level deals with pixel level features, high-level deals with abstract concepts and intermediate-level deals with something in between. Whilst low-level vision is fairly well studied and we have a good understanding at this level, middle and high level concepts are very difficult to grasp, certainly extremely difficult to represent using computer bits. In the signal processing community, an image can be represented in the time/spatial domain and in the frequency/spectral domain. Both time domain and frequency domain analysis technologies are very well developed [Akan92, Robe87]. A signal/image can be represented as time sequence or transform coefficients of various types - Fourier, Wavelet, Gabor, KLT, etc. These coefficients often provide a convenient way to interpret and exploit the physical properties of the original signal. Exploiting well-established signal and analysis technology to represent and interpret vision concepts could be a fertile area for making progress. The human vision system is extremely sophisticated and more
powerful. The human vision theories could provide guidance for building practical engineering solution to vision tasks [Mojs00].

Content-based image and vision indexing and retrieval have been a popular research subject in many fields related to computer science for over a decade. Of all the challenging issues, associated tasks with the indexing and retrieval tasks are probably most difficult to achieve. The difficulties can be explained from a number of perspectives. First, relevance is a high-level concept and is therefore difficult to describe numerically using computer bits. Secondly, traditional indexing approaches mostly extract low-level features in a low-level fashion and it therefore difficult to represent relevance using low-level features. Because low-level features bear no correlation to high-level concepts, the burden of relevant retrieval has to be on high-level retrieval strategies, which is again hard. We believe one way in which one can make progress is to develop numerical representations (low-level features) that not only have clear physical meanings but also be related to high-level perceptual concepts. From an engineering point of view, such representation should be easy to compute, efficient to store and which should also render simple and effective retrieval.

In this chapter, the similarity of the images is examined by comparing the variation among the pixels within and between the images, and the spectrum of energy between the query and target images. To test the variations among the pixels within and between the images, the F-ratio test is adopted, and to test the spectrum of energy between the query and target images, the Welch’s t-test, i.e. test for equality of two means is employed. If either the query or the target image is treated as a sample, the other
is treated as population. By testing the query and target images, it is inferred that whether there is variation and interaction among the pixels within and between the images. If the F-ratio test and the Welch’s t-test pass, then it is concluded that the two images are same or similar; otherwise, it is assumed that they differ.

2.2 TESTS FOR SIMILARITY OF IMAGES

Generally, when capturing an image through camera (digital or analog) or scanning through scanner, there are many possibilities to include noise in the image. The inclusion of noise is a random process, which is independent and identically distributed Gaussian random variable. As discussed in Chapter 1, let $X$ be a random variable that represents the intensity value with additive noise of a pixel at location $(k, l)$ in a colour image. Hence, an image is assumed to be Gaussian Markov Random field [Kris07, Seet07]. The pixel $X(k, l) \in \mathbb{R}^3$ is a linear combination of three colours, namely, red, green, and blue, i.e. $X(k, l) = \{r(k, l), g(k, l), b(k, l)\}$. The mean intensity value of each colour is represented by $\mu$, and the variation is denoted by $\sigma^2$.

The normal density function of $X(k, l)$ for each colour is given by

$$
\frac{1}{\sigma \sqrt{2\pi}} \exp \left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \ -\infty < x, \ \mu < \infty, \ \sigma > 0 \quad \text{... (2.1)}
$$

The density function in equation (2.1) can be denoted as,

$N(x/\mu, \sigma^2)$ and the distribution law as $N(\mu, \sigma^2)$.

2.2.1 Test Statistic for Interaction among the Pixels within and Between Images

The F-ratio test [Rand12][Ruxt06] is used to test the equality of variations or interactions among the pixels in the query and target images. Let the intensity values $F_q$ =
(f_q^1, f_q^2, ..., f_q^n) and Ft = (f_t^1, f_t^2, ..., f_t^n) of the query and target images be independent and identically distributed samples from two normal populations. In order to test the variations or interactions among the pixels within query or target image(s), and variations between the query and target images, the F-ration test is performed. The test statistic is expressed as in Equation (2.2).

The test of the hypothesis is framed for two-tailed test is as follows.

Hypotheses:

H_0: \sigma_q = \sigma_t (Similarity)

H_a: \sigma_q \neq \sigma_t (Non-similarity)

\[ F = \frac{\sigma_q^2}{\sigma_t^2} \quad \ldots \quad (2.2) \]

Let

\[ \sigma_q^2 = \frac{1}{n_q - 1} \sum_{i=1}^{n_q} (f_q^i - \bar{f}_q)^2 \quad \ldots \quad (2.3) \]

\[ \sigma_t^2 = \frac{1}{n_t - 1} \sum_{i=1}^{n_t} (f_t^i - \bar{f}_t)^2 \quad \ldots \quad (2.4) \]

be the sample variances of the query and target images respectively. Let

\[ \bar{f}_q = \frac{1}{n_q} \sum_{i=1}^{n_q} f_q^i \quad \ldots \quad (2.6) \]

and

\[ \bar{f}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} f_t^i \quad \ldots \quad (2.6) \]

be the sample means of the query and target images respectively; n_q and n_t are the
number of pixels in the query and target images respectively; \( f_q \) and \( f_t \) are the standard deviations of the query and target images.

**Critical Region:** If the F-ratio value is less than the critical value, \( F_{\alpha, n_q-1, n_t-1} \), with degrees of freedom \( n_q-1 \) and \( n_t-1 \), and the significance level \( \alpha \), it is inferred that the query and target images are same or similar; otherwise, it is concluded that the two images differ. The critical value \( F_{\alpha} \) is referred to the F distribution in the statistical table.

### 2.2.2 Test for Equality of Spectrum of Energy between the Query and Target Images

As discussed in the previous section, if the interaction among the pixels’ intensity values between the query and target images passes the test, then it is proceeded to test the equality of spectrum of energy between the query and target images with the basic assumption that there is no variation among the intensity values between the query and target images. After observing the outcome of the two tests, it is concluded that the two images are same or not. To achieve this, the test for equality of means, i.e. Welch’s t-test is performed on the images. The tests of hypotheses are assumed to be as follows:

If the test for equality of variances and the means manifest that the query and target images pass the tests, such as F-ratio and Welch’s t-test, it is concluded that the two images are same or similar; otherwise, it is assumed that they belongs to different groups. The Welch’s test statistics is expressed as in equation (2.7), which compares the mean values of the query and target images. The test of the hypothesis is framed for two-tailed test is as follows.
H₀: μᵣ = μₜ (Similarity)

Hₐ: μᵣ ≠ μₜ (Non-similarity)

\[ W_t = \frac{\mu_r - \mu_t}{\sqrt{\frac{s^2_r}{n_r} + \frac{s^2_t}{n_t}}} \] ... (2.7)

The μᵣ and μₜ are mean values of the query and target images that can be computed using the expression presented in equations (2.5) and (2.6) respectively. The \( s^2_r \) and \( s^2_t \) are the sample variances of the query and target images respectively those can be computed using the expressions in equations (2.3) and (2.4). The variances of the query and target images are computed individually separately. The degrees of freedom, ν, is calculated using the formula given in equation (2.8), which is used to refer to the statistical table for the significant level α.

\[ \nu = \frac{\left( \frac{s^2_r}{n_r} + \frac{s^2_t}{n_t} \right)^2}{\left( \frac{s^2_r}{n_r} \right)^2 + \left( \frac{s^2_t}{n_t} \right)^2} \] ... (2.8)

**Critical Region:** If the WT value is less than the critical value, \( t_{(\alpha, \nu)} \), with degrees of freedom, ν, and the significance level α, it is inferred that the query and target images belongs to the same class. Otherwise, it is concluded that they belongs to different classes of image datasets. The critical value \( t_\alpha \) is referred to as the t-distribution in the statistical table. The implementation of these two statistical tests are discussed as detailed in Chapter 5.